



ADMIRe DMP Online analysis

Author(s):	Mark Berry & Thomas Parsons
Audience:	Project stakeholders & JISC MRD
Published:	15/10/2012

Contents

1.	Introduction	2
2.	DMP Online	2
2.1.	3 Iterations of the plan.....	2
2.2.	Customising DMP Online.....	3
3.	Sections of DMP Online	5
4.	Further work	18

1. Introduction

The following report documents analysis carried out by the ADMIRe project with regards to using DMP Online. The analysis reviewed how DMP Online could be used and identified how it could best be integrated with existing Nottingham (The University of Nottingham) software.

2. DMP Online

[DMP Online](#) is an online tool produced by the DCC for building and editing DMPs according to the requirements stipulated by the major UK funders. Instructions for use of DMP Online are [here](#). The normal scenario is for the research team to answer the questions in DCC's hosted instance of DMP Online; the resulting plan can be exported as PDF, HTML or DOCX at any stage, with the sections re-ordered and included/excluded as desired.

This document reproduces the [DMP Checklist v 3.0 \(DCC\)](#) - a table of all the DMP Online Questions and default guidance - and indicates which sections overlap with metadata held in other Nottingham business systems (principally Agresso and the proposed EQUELLA¹ metadata record) and which sections are potentially customisable by the Nottingham and/or schools as templates.

2.1. 3 Iterations of the plan

The DMP Online plan is designed to be completed in 3 iterations:

1. **Minimal Plan:** For use at the grant application stage. When creating a plan, you select which funders or other organisations to base the plan on, and this activates templates to include only the subset of questions required by those funders at the application stage. By answering these questions, the funders' requirements should be met.
2. **Core Version:** Developed at early-project stage and maintained throughout the project lifecycle. Once funding is in place, change the basic project information from 'unfunded' to 'funded' and a much longer series of questions are presented, covering all aspects of in-project data management. These do not all have to be answered: additional questions can be added and unwanted questions can be hidden as required.
3. **Full Plan:** Addresses issues of long-term preservation and access. This adds Section 6 of the plan (long-term preservation), several questions in Section 1 (basic project information), and Sections 9 and 10 (statement of agreement and annexes). Elements of the Full Plan which are not present in the Core Plan are indicated below as 'NOT CORE'.

¹ <https://test.equella.nottingham.ac.uk/institutions.do>

2.2. Customising DMP Online

As explained by [Customising DMP Online for your institution](#), there are 3 options for institutional customisation of DMP Online:

1. **Customised Institutional Template** - Select subsets of the DCC Checklist Questions and add guidance tailored to the institution (including web and email links to institutional support) and boilerplate responses to some questions. Send the template to the DCC to add it to the DMP Online system. The institution's logo can appear within DMP Online while users are entering their answers.
2. **Custom-Branded Institutional Version** - More control over the tool's appearance is possible by customising style sheets and using a custom URL (e.g. dmponline.nottingham.ac.uk). This requires institutional commitment: a named administrator responsible for on-going template and style sheet maintenance and support from IS to enable the custom URL.
3. **Self-Hosted Instance** - The open-source software can be hosted by the institution, but this entails responsibility for the upkeep of the system and, in particular, changes to funder requirements, checklist questions and guidance.

The two tables of contents below set out the DMP structure in outline and in detail. Below these tables of contents, the full guidance from the standard DMP Online Templates can be seen for each question.

IMPORTANT NOTES

1. **EQUELLA Metadata:** Where questions overlap with metadata from EQUELLA, this is indicated in green, and where sections contain such questions this is also indicated in the section heading. Note that, depending on the workflow, it may be possible to populate the data for these questions from EQUELLA, or (more likely) the reverse may be more appropriate: populate the metadata in EQUELLA from the DMP. A mechanism for copying this content from one system to the other remains to be determined, and will probably require either the provision of an API for DMP Online, or the provision of XML output from DMP Online to EQUELLA. Full integration of the two systems may even require that the Nottingham uses the Self-Hosted Instance option.

2. **Nottingham Templates:** The questions marked in red as 'Nottingham Template' are all the questions which can **potentially** be customised with standard Nottingham Template content. With most of these questions, the research team will probably still need to enter further detail. **All sections that are not marked as Nottingham Template are not capable of such standardisation and must be completed by the research team.** In some cases, School-level standard content may also be appropriate, and this is indicated in red using "Nottingham+ SCHOOL: Template", but any school-level content will probably have to be provided directly to researchers by their schools rather than as templates within DCC Online. Nottingham may also wish to customise the template **advice** (beyond that offered by the DCC in the default templates) for all the other questions as well - see 'Website Content' below for more on this.
3. **Website Content:** The bulk of the DMP Questions overlap considerably with the sections and content of the proposed Nottingham Research Data Management Website². It will therefore be desirable to harmonise the content of the website with the content of the DMP Online templates, and it is recommended that they should be designed in parallel. The DMP Online templates could all link to the appropriate page of the Nottingham RDM Website for further guidance, and the Website pages should probably be structured according to the DMP Online Questions. This applies to nearly all sections of DMP Online.
4. **Projects without funders' requirements for DMP:** DMP Online is based on the choice of funder, which dictates the minimal and core versions of the plan. For projects without funders that mandate such requirements, a separate institutional template may be necessary. Projects still need data management plans even if funders don't make this a requirement. Finding a generic model for projects without funders' requirements may prove challenging. Based on Northampton's experience, it seems likely that this will have to be implemented as a recommendation rather than as a requirement.
5. **Bath's DMP Template for Engineering also adds the following useful sections**, with detail not specifically covered in DMP Online – these sections should be considered for inclusion in the Nottingham DMP Online Templates:
 - Link to Project Record Manifest
 - Plan for data generation and manipulation
 - Plan for data organisation (file naming convention, version control, full metadata)
 - Plan for data quality
 - Details of data structures and formats
 - Explanation of data semantics

² <http://www.nottingham.ac.uk/research/research-data-management/index.aspx>

3. Sections of DMP Online

The following sections document what is found within DMPOnline. They identify where interactions with other Nottingham systems would be beneficial.

1.0 Introduction and Context

This section records administrative details which tie the plan to a particular project.

1.1 Basic Project Information **SYSTEM CALLS: Agresso/EQUELLA**

1.1.1 Project name **SYSTEM CALL: Agresso/EQUELLA**

1.1.2 Funding body (or bodies) (NOT CORE) **SYSTEM CALL: Agresso/EQUELLA**

1.1.3 Budget (NOT CORE) **SYSTEM CALL: Agresso**

1.1.4 Duration (NOT CORE) **SYSTEM CALL: Agresso/EQUELLA**

1.1.5 Lead partner organisation (NOT CORE) **SYSTEM CALL: Agresso**

1.1.6 Other partner organisations (NOT CORE) **SYSTEM CALL: Agresso**

1.2 Short description of the project's fundamental aims and purpose **SYSTEM CALL: Agresso/EQUELLA**

Information summarised from the main body of your research proposal will help potential re-users understand the purposes your data has been collected or created for, and they are unlikely to have access to your proposal. Briefly summarise what you set out to discover and how that is likely to affect the kind of data you collect or create and how.

1.3 Related Policies **SYSTEM CALL: JULIET; Nottingham+SCHOOL: Templates**

Some of the information you give in the remainder of the DMP will be determined by the content of other policies; these policies may also have additional requirements that are not covered here. In case of doubt it is helpful for data managers to know what other policies were in force when the DMP was written.

1.3.1 Funding body requirements relating to the creation of a data management plan **SYSTEM CALL: JULIET**

Guidance:

[DCC comparison of Research Funders' DMP Requirements](#)

1.3.2 Institutional or research group guidelines **Nottingham+SCHOOL: Template(s)**

For multi-partner projects, you may also wish to mention any formal consortium agreement agreed, e.g. on data sharing, publication, IPR.

1.3.3 Other policy-related dependencies **Nottingham+SCHOOL: Template**

Examples of other relevant policies may include institutional ethics, regulation, information governance, and guidance and requirements from the data centre to which the data will be submitted.

1.4 Basic Data Management Plan Information **SYSTEM: Automatic; Nottingham+SCHOOL: Templates**

1.4.1 Date of creation of this plan **SYSTEM: Automatic**

Recording date information is important for version control and placing the DMP in context.

1.4.2 Aims and purpose of this plan **Nottingham+SCHOOL: Template**

Here you may wish to address the following: protecting IPR, protection of sensitive data, adding value, ensuring longer term access, etc.

Target audience for this plan **Nottingham+SCHOOL: Template**

2.0 Data Types, Formats, Standards and Capture Methods

It is of critical importance that research datasets are adequately documented. The information in this section will help you and any subsequent user understand why and how the data were created, what they represent, and whether they are likely to be compatible with other datasets.

2.1 Give a short description of the data being generated or reused in this research

When describing the type of content to be created, you may wish to refer to the RIN data types as a way of classifying what you will create: Scientific experiments; Models or simulations; Observations; Raw data; Derived data; Canonical or reference data. (See "To Share or not to Share: Publication and Quality Assurance of Research Data Outputs", Research Information Network, 2008)³.

You should also consider the implications of data volumes: do you have sufficient storage? Will the scale of the data pose challenges when sharing or transferring data between sites?

2.2 Existing Data

2.2.1 Have you reviewed existing data, in your own institution and from third parties, to confirm that new data creation is necessary? (YES/NO)

Performing this check helps also helps to ensure the uniqueness of the research.

2.2.2 What existing datasets could you use or build upon?

If none, enter "n/a"

2.2.3 Describe any access issues pertaining to the pertinent, existing data

If relevant, include financial costs of accessing or using the data.

2.3 New Data **SYSTEM CALL: EQUJELLA; Nottingham: Template**

2.3.1 Why do you need to capture/create new data?

Reasons to capture/create new data will include: non-existence of suitable existing data; extending existing data to cover new areas; performing comparison over time.

2.3.2 Describe the process by which you will capture/create new data

³ <http://www.rin.ac.uk/our-work/data-management-and-curation/share-or-not-share-research-data-outputs>

Here you should explain the capture process. If you're doing observations, how will they be recorded? (e.g. in a dated and numbered field notebook.) Also note what kind of equipment you will use and the software required. If you plan to use proprietary software, could you export to an Open format so the data can be reused more widely? You may also wish to cover: content selection; instrumentation; technologies and approaches chosen; file naming conventions; versioning; meeting user needs. Your answer should be sensitive to the location in which data capture will take place. Guidance:

JISC digital media guidance on file-naming:

<http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/choosing-a-file-name/>

University of Edinburgh Records Management file naming guidance:

<http://www.recordsmanagement.ed.ac.uk/InfoStaff/RMstaff/RMprojects/P/FileNameRules/Rules.htm>

2.3.3 Which file formats will you use, and why? **SYSTEM CALL: EQUELLA**

Here you should outline and justify your choice of format, e.g. Microsoft Excel for recording measurements or SPSS for analysis, as these are in widespread use, the University has the relevant software licences or they're accepted standards in your field, etc. Decisions relating to file formats may also be made with recourse to staff expertise, a preference for Open formats, accepted standards, or widespread usage with a given community. Guidance:

UKDA Guidance on recommended data formats:

<http://www.data-archive.ac.uk/sharing/acceptable.asp>

2.3.4 What criteria will you use for Quality Assurance/Management? **Nottingham: Template**

Quality management mechanisms may include: documentation, calibration, validation, monitoring, transcription metadata, peer-review.

2.4 Relationship between old and new data

2.4.1 What is the relationship between the new dataset(s) and existing data?

This is concerned less with existing data that may be used in the Research Activity, but rather with the disciplinary context. A typical answer might identify a body of data with which it would be helpful to harmonise newly generated data, or from which methodologies might be drawn, e.g. ISO standard materials testing data, time/motion studies data.

2.4.2 How will you manage integration between the data being gathered in the project and pre-existing data sources?

Here you may wish to cover issues such as technical integration, provenance, trust and data quality.

2.4.3 What added value will the new data provide to existing datasets?

Value which new data can bring to old may include: greater detail, wider coverage, verification of existing data, etc.

2.5 Data Documentation and Metadata **Nottingham+School: Template**

Metadata is the information that makes your new data usable. NISO defines three main categories of metadata: Descriptive metadata is the information used to search and locate an object such as title, author, subjects, keywords, publisher; structural metadata gives a description of how the components of the object are organised; and administrative metadata refers to the technical information including file type. Two sub-types of administrative metadata are rights management metadata and preservation metadata. (Source: Wikipedia)

Annotation briefing paper:

<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/annotation>

2.5.1 Are the datasets which you will be capturing/creating self-explanatory, or understandable in isolation? (YES/NO)

You may wish to consider this from the perspective of a typical reader of a journal for your discipline.

2.5.2 If you answered No to DCC 2.5.1, what contextual details are needed to make the data you capture or collect meaningful?

Think about what kind of documentation is needed for others to understand your data. This may include: a description of the data capture methods, explanation of data analysis, details of who has worked on the project and performed each task, etc.
Guidance:

JISC Digital Media Introduction to Metadata:

<http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/an-introduction-to-metadata/>

UKDA Guidance on Data Documentation and Metadata:

<http://www.data-archive.ac.uk/sharing/metadata.asp>

2.5.3 How will you create or capture these metadata? **Nottingham+School: Template**

You may wish to address the balance between automatic and manually created metadata. Creating documentation takes time so consider whether anything you're already creating can be used e.g. publications, websites, progress reports, etc. Also note where information about the data will be recorded e.g. in a database with links to each item, in a 'readme' text file, in file headers / under properties in Word or PDF.
Guidance:

DCC Briefing Paper on Annotation:

<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/annotation>

2.5.4 What form will the metadata take? **Nottingham+School: Template**

Where appropriate, give details of the standards used. Using standards such as Dublin Core and TEI can make your data interoperable, so consider what others in your field have used or follow data centre recommendations. Using controlled vocabularies for description will also help improve consistency. Guidance:

DCC Briefing Paper on Metadata Standards:

<http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/what-are-metadata-standards>

2.5.5 Why have you chosen particular standards and approaches for metadata and contextual documentation? **Nottingham+School: Template**

Decisions relating to metadata standards may be made with recourse to: staff expertise, a preference for Open standards, or widespread usage with a given community. Guidance:

DCC Briefing Paper on Metadata Standards:

<http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/what-are-metadata-standards>

3.0 Ethics and Intellectual Property

Certain types of data impose additional ethical and legal constraints on how data should be used and managed. Data use can be hampered by a lack of clarity over intellectual property rights.

3.1 Ethical and Privacy Issues **SYSTEM CALL: EQUELLA**

3.1.1 Are there ethical and privacy issues that may prohibit sharing some or all of the dataset(s)? (YES/NO); **SYSTEM CALL: EQUELLA**

Guidance:

UKDA Guidance on Consent, Confidentiality and Ethics:

<http://www.data-archive.ac.uk/sharing/confidential.asp>

3.1.2 If you answered Yes to DCC 3.1.1, How will these be resolved? **SYSTEM CALL: EQUELLA**

Ways to resolve these may include: anonymisation of data; referral to departmental or institutional ethics committees; or formal consent agreements. The consent agreements you make with research participants and Data Protection legislation affect how you store data, who can see/use it and how long it is kept. You should show that you're aware of this and have planned accordingly.

3.1.3 Is the data that you will be capturing/creating "personal data" in terms of the Data Protection Act (1998) or equivalent legislation if outside the UK?

Guidance:

DCC Briefing Paper on Data Protection:

<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/data-protection>

3.1.4 What action will you take to comply with your obligation under the Data Protection Act (1998) or equivalent legislation if outside the UK?

3.2 Intellectual Property Rights **SYSTEM CALL: EQUELLA**

It is important to strike an appropriate balance between concern for legal implications and getting research done. Inactivity due to legal overwhelm is better avoided!

3.2.1 Will the dataset(s) be covered by copyright or the Database Right? If so give details in DCC 3.2.2, below. (YES/NO); **SYSTEM CALL: EQUELLA**

Guidance:

DCC Legal Watch Paper on the Database Right:

<http://www.dcc.ac.uk/resources/briefing-papers/legal-watch-papers/ipr-databases>

3.2.2 If you answered Yes to DCC 3.2.1, Who owns the copyright and other Intellectual Property? **SYSTEM CALL: EQUELLA**

For multi-partner projects, this may be worth covering in a consortium agreement. Ideally, this should address the risk of movement of staff between institutions mid-project.

3.2.3 If you answered Yes to DCC 3.2.1, How will the dataset be licensed? **SYSTEM CALL: EQUELLA**

Any restrictions on use should be justified, and a timeframe for data release outlined to assure the funder of wider public benefit where possible. For example will there be: delays in releasing data while you seek a patent? Planned embargo periods / right of first use to secure publications? Prevention of data sharing due to terms of commercial partnership agreements?

Guidance:

DCC Legal Watch Paper on Creative Commons:

<http://www.dcc.ac.uk/resources/briefing-papers/legal-watch-papers/creative-commons-licensing>

DCC Legal Watch Paper on Science Commons:

<http://www.dcc.ac.uk/resources/briefing-papers/legal-watch-papers/science-commons>

3.2.4 For multi-partner projects, what is the dispute resolution process / mechanism for mediation?

You may wish to cover this in a consortium agreement, in which case you can just answer "As per the consortium agreement."

4.0 Access, Data Sharing and Reuse

There are often conflicting pressures on researchers to share or withhold their data. Early consideration of the issues can help to resolve these conflicts.

4.1 Access and Data Sharing **Nottingham: Template**

4.1.1 Are you under obligation or do you have plans to share all or part of the data you create/capture? (YES/NO)

Your funding body may insist on data sharing, and - if you are in the UK - your project may be subject to Freedom of Information (FoI) legislation. (Note that FoI legislation differs in Scotland from England and Wales.)

Guidance:

UKDA Guidance on Data Sharing:

<http://www.data-archive.ac.uk/sharing/whyshare.asp>

4.1.2 If you answered No to DCC 4.1.1, why will you not share your data?

You may not plan to share data due to: ethical reasons; non-disclosure agreements; or quality-related issues. (You may also choose to share only part of your dataset(s): if so, give details here.)

Guidance:

DCC Legal Watch Paper on Sharing Medical Data:

<http://www.dcc.ac.uk/resources/briefing-papers/legal-watch-papers/sharing-medical-data>

4.1.3 If you answered Yes to DCC 4.1.1, How will you make the data available?
Nottingham: Template

Here you will want to explain how the data will be shared e.g. will they be deposited in a data centre, will you forward copies on request to interested parties, etc. Also consider how potential users will find out about your data, e.g. will you publish details of your research, present at conferences, blog about your findings, promote your research outputs on a website? etc.

4.1.4 If you answered Yes to DCC 4.1.1, When will you make the data available?

4.1.5 If you answered Yes to DCC 4.1.1, What is the process for gaining access to the data?

Ways of accessing data include: downloading from a data centre; requesting direct from the researcher; downloading from a Web page.

4.1.6 If you answered Yes to DCC 4.1.1, Will access be chargeable? (YES/NO)

4.1.7 If you answered Yes to DCC 4.1.6, Please give details.

4.2 Exploitation **SYSTEM CALLS: EQUELLA**

Exploitation of data may comprise using the data in support of academic publications, or for some other kind of gain (e.g. commercial).

4.2.1 Does the original data collector/ creator/ principal investigator retain the right to use the data before opening it up to wider use? (YES/NO)

4.2.2 If you answered Yes to DCC 4.2.1, Please give details. **SYSTEM CALL: EQUELLA**

All the funders that we've examined permit embargoes, but expect them to be reasonable and expect justification (e.g. for the time limits set).

4.2.3 Are there any embargo periods for political/commercial/patent reasons? (YES/NO) **SYSTEM CALL EQUELLA**

4.2.4 If you answered Yes to DCC 4.2.3, Please give details. **SYSTEM CALL: EQUELLA**

4.3 Reuse **SYSTEM CALL: EQUELLA**

4.3.1 Which groups or organisations are likely to be interested in the data that you will create/capture? **SYSTEM CALL: EQUELLA**

There is a push for publicly funded data to be of wide benefit, so it may help to show that you envisage your data being of use beyond your group, or even beyond your discipline.

4.3.2 How do you anticipate your new data being reused?

Explain how the data will be developed with future users in mind, i.e. are your choices of formats, technologies and metadata appropriate to these audiences?

5.0 Short-Term Storage and Data Management

You should note what support is provided, e.g. "we will use the University's networked service, which is backed up daily by computing support." Or, if you will manage your own storage and backup, explain how you will do that, noting any agreements you have in place e.g. mirroring data on a second server at the project partner's University. Additionally, more and more researchers keep data on portable devices (laptops, USB sticks, etc). It is crucial that short-term storage policies address and make provision against unintended loss of portable equipment.

5.1 Storage Media and Data Transfer **Nottingham: Templates**

This section relates primarily to in-project storage, as opposed to longer-term storage/preservation.

5.1.1 Where (physically) will you store the data during the project's lifetime? **Nottingham: Template**

Storing data on laptops alone is very risky: backed-up network drives are far preferable.

Guidance:

UKDA Guidance on Data Storage:

<http://www.data-archive.ac.uk/sharing/datastorage.asp>

5.1.2 What media will you use for primary storage during the project's lifetime? **Nottingham: Template**

5.1.3 How will you transfer/transmit the data, if this is required? **Nottingham: Template**

You may need to consider the data transfer speeds supported by your primary storage device, and if possible seek guidance from your institution's computing service on whether the available bandwidth on the local network, and your institution's network infrastructure, will be sufficient to meet your project's needs for short term collaborative working and any Web-based data publication. You may also want to address encryption if this is appropriate/necessary, and whether it is appropriate to transfer your data across unsecured network connections.

5.2 Back-Up **Nottingham: Templates**

5.2.1 How will you back-up the data during the project's lifetime? **Nottingham: Template**

Remember to consider all of the costs of backup, e.g. logging storage locations, version control, and of recovering data from the backup. These time/staff costs will far exceed the price of the storage device. If these are set against the risks of the device failing, becoming lost, destroyed or unusable, a centralised backup service is more likely to be justifiable. This service may be provided by your institution; you may also choose to incorporate off-site storage for additional protection, or arrange your own backup regime.

Guidance:

UKDA Guidance on Data Backup:

<http://www.data-archive.ac.uk/sharing/backup.asp>

5.2.2 How regularly will back-ups be made? **Nottingham: Template**

This may be something you choose to leave to your institutional or departmental support, but it's worth recording the information here.

5.2.3 Who is responsible for backup? **Nottingham: Template**

5.3 Security **Nottingham: Templates**

Security decisions may be made with a view to your data's financial value and/or its sensitivity.

5.3.1 How will you manage access restrictions and data security during the project's lifetime? **Nottingham: Template**

This may be managed via various levels of password protection.
Guidance:

DCC Briefing Paper on Information Security Management:

<http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/information-security-management-iso-27000-iso-27k-s>

UKDA Guidance on Data Security:

<http://www.data-archive.ac.uk/sharing/security.asp>

5.3.2 How will you implement permissions, restrictions and/or embargoes? **Nottingham: Template**

You may wish to give details of any policies in place governing making copies of data.

5.3.3 Give details of any other security issues. **Nottingham: Template**

6.0 Deposit and Long-Term Preservation (NOT CORE)

Section 6 is about long-term preservation. Many researchers will not perform these tasks themselves, so data centre staff or other long-term stewards may be best placed to answer these questions.

Guidance:

DCC Briefing Paper on Digital Repositories:

<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/digital-repositories>

6.1 What is the long-term strategy for maintaining, curating and archiving the data? (NOT CORE)

Nottingham: Template

Here you will want to demonstrate consultation between data creators and the relevant repositories / data centres to secure an appropriate place of deposit. Give details on the rationale for choosing this particular place of deposit. (N.B. Funders may require data to be offered to a particular data centre on completion of the project.) If there isn't anywhere you can deposit, explain how you will address sustainability e.g. by choosing open standards, or note how your institution can support you to store and manage the data in the longer term. Remember that you can consult institutional archivist(s) and records managers in formulating long-term retention plans.

Guidance:

DCC Briefing Paper on Digital Preservation:

<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/digital-preservation>

JISC Briefing Paper on Digital Preservation:

<http://www.jisc.ac.uk/media/documents/publications/digitalpreservationb.p.pdf>

6.2 Long-Term Specifics (NOT CORE) Nottingham: Templates

This section addresses three key issues: Selection, Retention, and Transformation.

6.2.1 Will or should data be kept beyond the life of the project? (YES/NO) (NOT CORE)

6.2.2 If you answered Yes to DCC 6.2.1, How long will or should data be kept beyond the life of the project? (NOT CORE) Nottingham: Template

Your funding body or institution may specify time-spans for retention. If not, general guidance is given in the RCUK Code of Good Research Conduct which says that "data should normally be preserved and accessible for ten years, but for projects of clinical or major social, environmental or heritage importance, for 20 years or longer."

6.2.3 If you answered Yes to DCC 6.2.1, What data centre/ repository/archive have you identified as the long-term place of deposit? (NOT CORE) Nottingham: Template

Your funder may have a preferred place of deposit.

6.2.4 What data will be preserved for the long-term? (NOT CORE)

You may wish to preserve all, none, or a selection of data over the long-term. You should also indicate here whether you will preserve raw data, derived data, samples, etc.

6.2.5 On what basis will data be selected for long-term preservation? (NOT CORE)

You may wish to include timeframes here as well.
Guidance:

DCC Briefing Paper on Appraisal and Selection:

<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/appraisal-and-selection>

6.2.6 If the dataset includes sensitive data, how will you manage this over the longer term? (NOT CORE) **Nottingham: Template**

This should include a justification of decisions and should cover deletion of data if appropriate.

6.2.7 Will transformations be necessary to prepare data for preservation and/or data sharing? (YES/NO) (NOT CORE)

Examples of transformation may include data cleaning/anonymisation where appropriate, or migration to another file format.

6.2.8 If you answered Yes to DCC 6.2.7, what transformations will be necessary to prepare data for preservation / future re-use? (NOT CORE)

Examples of transformation may include data cleaning/anonymisation where appropriate, or migration to another file format.

6.3 Metadata and Documentation for Long-Term Preservation (NOT CORE)
Nottingham: Templates

If you are a researcher submitting your data to a data centre or repository, the earlier you consider their metadata and documentation requirements the less painful it will be to provide the essential details, the better the chances of your data being found and re-used, and therefore the higher the chance of it having a lasting impact. Here you will want to show that you are aware of data centre standards for deposit, and have reflected these in your data development plans. You may wish to include (e.g.) references, reports, research papers, fonts, the original bid proposal, etc. You may also wish to include contextual/ related/ representation information.

6.3.1 What metadata/ documentation will be submitted alongside the datasets or created on deposit/ transformation in order to make the data reusable? (NOT CORE) **Nottingham: Template**

Digital files are fundamentally strings of binary digits (bits). In order to process them, one must know the format they are in and what software is needed to read that format. Even after the file has been successfully opened, extra information may be needed in order to fully understand the contents. In the terms of the Open Archival Information System (OAIS) Reference Model, the information required to transform a stream of bits into something intelligible is called representation information.
Guidance:

DCC Glossary Definition of Representation Information:

<http://www.dcc.ac.uk/digital-curation/glossary>

6.3.2 How will this metadata/documentation be created, and by whom? (NOT CORE) **Nottingham: Template**

The AHDS Catalogue Form is used to produce a full catalogue record for online catalogues.

Guidance:

AHDS Catalogue Form:

<http://www.ahds.ac.uk/depositing/catalogue-form.htm>

6.3.3 Will you include links to published materials and/or outcomes? (YES/NO) (NOT CORE)

6.3.4 If you answered Yes to DCC 6.3.3, please give details. (NOT CORE)

6.3.5 How will you address the issue of persistent citation? (NOT CORE)
Nottingham: Template

You may wish to refer to Digital Object Identifiers (DOIs), Persistent URLs, etc.

Guidance:

DCC Briefing Paper on Persistent Identifiers:

<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/persistent-identifiers>

The Digital Object Identifier System:

<http://www.doi.org/>

6.4 Longer-Term Stewardship (NOT CORE) **Nottingham: Templates**

6.4.1 Who will have responsibility over time for decisions about the data once the original personnel have gone? (NOT CORE) **Nottingham: Template**

This is likely to be either an institutional library or repository, or some other data custodian (e.g. a data centre.)

6.4.2 In the event of the long-term place of deposit closing, what is the formal process for transferring responsibility for the data? (NOT CORE) **Nottingham: Template**

This should be completed by a representative of the original place of deposit.

7.0 Resourcing

It is important that data management is treated as a first-class research activity, with appropriate funds and effort allocated to it.

7.1 Outline the staff/organisational roles and responsibilities for implementing this data management plan. **Nottingham+School: Template**

This could include: data management time allocations; project management of technical aspects; training requirements; storage and backup; contributions of non-project staff, etc. Individuals should be named where possible. Continue in an Annex if necessary.

7.2 How will data management activities be funded during the project's lifetime?

This should cover (e.g.) payments to service providers within institutions, payments to external data centres for hosting data, income derived from licensing data, etc). It is also important to remember to build costs of in-project data management into the project budget. (N.B. Some funders state explicitly that they will meet the cost of preparing data for deposit, so remember to include this in your time and budget allocation too!)

7.3 How will longer-term data management activities be funded after the project ends? (NOT CORE) **Nottingham: Template**

This should cover (e.g.) payments to service providers within institutions, payments to external data centres for hosting data, income derived from licensing data, etc). It is also important to remember to build costs of in-project data management into the project budget.

8.0 Adherence and Review

A data management plan can only be effective if everyone agrees to adhere to it. Communication is also important: the human aspects of data management are widely held to be more difficult than the technical aspects. All aspects of the environment will change over time, so it is similarly important that you keep your plans up-to-date via regularly scheduled review.

8.1 Adherence [Nottingham: Templates](#)

8.1.1 How will adherence to this data management plan be checked or demonstrated? [Nottingham: Template](#)

8.1.2 Who will check this adherence? [Nottingham: Template](#)

8.2 Review [Nottingham: Templates](#)

A data management plan should be a living document, and it is important - especially for longer term work - that it is reviewed on an ongoing basis.

8.2.1 When will this data management plan be reviewed? [Nottingham: Template](#)

8.2.2 Who will carry out reviews? [Nottingham: Template](#)

8.2.3 Does this version of the DMP supersede an earlier plan? (YES/NO)

8.2.4 If you answered Yes to DCC 8.2.3, you may wish to enter information about the relationship between versions here.

You may want to consider including previous versions of the plan as Annexes to this version.

9.0 Statement of Agreement (NOT CORE)

You may wish to formalise your Data Management Plan with a statement of agreement, including signatories if relevant.

9.1 Statement of Agreement (NOT CORE) [Nottingham: Template](#)

10.0 Annexes

10.1 Contact details and expertise of nominated data managers / named individuals [Nottingham: Template](#)

10.2 Glossary of terms (NOT CORE)

10.3 Other annexes as required (NOT CORE)

You may wish to attach other Annexes to your data management plan. You can use this space to list these Annexes.

4. Further work

The review was designed to promote discussion and identify where DMP Online could be used and what additional information would be needed. It has clearly indicated the complexity of the tasks facing researchers and there need for dedicated RDM support to aid them to create DMPs.

This information has been utilised and it is expected that Nottingham will host an instance of DMP Online from 2014 onwards, once the metadata and technical systems have been developed.